Model-based identification of chemicals transformation pathways combined with reaction kinetics models– the case of heroin biomarkers in wastewater

Ramin, Pedram; Valverde Pérez, Borja; Polesel, Fabio; Gernaey, Krist; Plósz, Benedek G.

Publication date:
2017

Document Version
Peer reviewed version

Frontiers International Conference on Wastewater TreatMent
21-24 MAY 2017 | PALERMO – ITALY

Model-based identification of chemicals transformation pathways combined with reaction kinetics models– the case of heroin biomarkers in wastewater

P. Ramin*', B. Valverde-Pérez *, F. Polesel*, K.V. Gernaey**, B. Gy. Plósz*'*****
* Department of Environmental Engineering, Technical University of Denmark, Lyngby, Denmark
** Department of Chemical and Biochemical Engineering, Technical University of Denmark, Lyngby, Denmark
*** Department of Chemical Engineering, University of Bath, Claverton Down, UK

Model identification; uncertainty propagation; illicit drug biomarkers

Introduction. Models are mathematical representations of real systems. To predict the behaviour of the real system with as small deviation as possible, model structure selection and model parameters estimation should be performed in a systematic procedure referred to as model identification. This task can be quite challenging for heterogeneous biological systems such as wastewater, which are usually described with rather complex mathematical formulations in order to properly represent the interactions between multiple types of microorganisms. A typical drawback of such models is over-parametrization, causing non-identifiability of model parameters. This issue can cause more problems when measurements are scarce and data quality is poor. Parameter identifiability issues can be tackled with effective experimental design combined with multi-faceted parameter estimation protocols, such as sensitivity analysis, parameter correlation, or iterative procedures (Brun et al., 2002; Mannina et al., 2011; Sin et al., 2005). However, these procedures are not reported for all biological processes and systems. For instance, despite increasing popularity of micropollutants fate models (e.g., Plósz et al., 2012; Polesel et al., 2016) in wastewater, the challenges related to model identification are less understood and addressed. The main objective of this study includes the development of a model-based identification method for multi-branched chemical transformation pathways combined with reaction kinetics models. The outcome can serve as a complementary tool for chemical pathway identification, which are commonly characterized by analytical chemistry methods. Heroin and codeine together with 4 other human metabolites in wastewater were considered as case study.

Methods. Transformation of heroin (HER) drug biomarkers and its human metabolites, 6-monoacetylmorphine (6-MAM), morphine (MOR), and morphine-3-β-D-glucuronide (MORG), (COE) and its human metabolite norcodeine (NCOE) in untreated wastewater under anaerobic conditions were used as case study. Specific details regarding the experimental assessment and process model development are presented by Ramin et al (2016). The model includes kinetic parameters, namely abiotic transformation rates, k_{abio} (d^{-1}), and biotransformation rate constants, k_{bio} (L g TSS d^{-1}), for each selected drug biomarker. Through model calibration against batch experimental, these parameters were estimated using the Bayesian optimization algorithm $DREAM_{(ZS)}$ (Laloy and Vrugt, 2012) following three different methods. Method 1 is the proposed method of choice, whilst the two other methods are commonly encountered in literature:

- *Method 1*: parameter estimation is performed by considering parameter ranges and associated probability distributions obtained at any given transformation pathway level as priors for parameter estimation at any subsequent transformation levels.
- *Method 2*: parameters are estimated in a concerted way and by omitting a priori information regarding the range and the parameter probability distribution – also referred to as the “lumped” estimation method.
- *Method 3*: Parameters are estimated using a step-wise approach (as in *Method 1*) by employing fixed parameter values via parameter estimation at downstream levels without considering the distribution of parameters.

Model accuracy predictions derived using the three different calibration methods were assessed using statistical tests. Values obtained for the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE) were related to model prediction accuracy, i.e. the lower the RMSE and MAE, the more accurate the model prediction. Additionally, the average relative Interval Length to Coverage of measurements by prediction bands (ILTC) was used for model prediction unertainty. A lower ILTC provides an indication of lower prediction uncertainty together with higher coverage of measured values by the uncertainty band.

Results. An overview of identification methodology (*Method 1*) is presented in Fig. 1a. The method involves three stages: (I) estimation of parameters at defined levels; (II) evaluation of model structure (i.e., kinetic and/or pathway) at each level; and (III) propagation of parameter uncertainty to the subsequent levels. The estimation methodology consists of n levels, defined according to assumed abiotic (A) and biotic (B) reaction pathways. The chemical pathway is represented by m chemicals (from X_1 to X_m) transformed via abiotic (e.g., hydrolysis) and microbially-mediated reactions. In the case of systematic error in the model prediction, the structure of the kinetic model (e.g., second-order or first-order kinetics) and the chemical transformation pathway (e.g., new transformation products) are re-evaluated (Fig. 1b). This iterative approach ensures that uncertainty of parameters would only propagate if the model structure is adequately describing the system. Following our previous study (Ramin et al., 2016) it is assumed that the structure of the model is sufficiently precise.
As to the transformation pathway of HER and its metabolites, the histograms from *Method 1* and *Method 3* (Fig. 2, red bars and solid black lines) show that for most parameters the posterior distributions are skewed (e.g., significant positive skewness for MOR (B1)). This is not the case for most of the parameter subsets obtained using *Method 2*, showing comparably higher values and predominantly wider uniform distribution. Following parameter estimation, the impact of parameter uncertainty on model outputs (drug concentrations) was assessed through uncertainty analysis using Monte Carlo simulations (Fig. 3). These results suggest that *Method 2* can lead to estimates with higher uncertainties for both abiotic and biotic model parameters. The prediction accuracy obtained using *Method 1* and *Method 3* (Table 1) showed comparable performance in the biotic model, except for MOR and 6-MAM, for which *Method 1* resulted in higher accuracy, and for MORG, where *Method 3* showed a better performance. Based on the collinearity threshold for identifiability (0.7;Sin et al., 2009)), all estimated parameters from all methods were identifiable except the parameters related to the chemicals with two transformation branches in *Method 3* (i.e., HER and MORG transformation pathways, Fig. 4).
Overall, these results indicate that *Method 1*, compared to the other two methods, can improve the identifiability of model parameters whilst maintaining a good prediction performance.

References

Brun, R., Siegrist, H., Gujer, W., Reichert, P., 2002. Practical identi ability of ASM2d parameters—systematic selection and tuning of parameter subsets. Water Res. 36, 4113–4127.
Laloy, E., Vrugt, J. a., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM(ZS) and high-performance computing. Water Resour. Res. 48, 1–18.
Mannina, G., Cosenza, A., Vanrolleghem, P. a., Viviani, G., 2011. A practical protocol for calibration of nutrient removal wastewater treatment models. J. Hydroinformatics 13, 575.
Plósz, B.G., Langford, K.H., Thomas, K. V., 2012. An activated sludge modeling framework for xenobiotic trace chemicals (ASM-X): Assessment of diclofenac and carbamazepine. Biotechnol. Bioeng. 109, 2757–2769.
Polesel, F., Andersen, H.R., Trapp, S., Plósz, B.G., 2016. Removal of Antibiotics in Biological Wastewater Treatment Systems □ A Critical Assessment Using the Activated Sludge Modeling Framework for Xenobiotics (ASM-X). Environ. Sci. Technol. 50, 10316−10334.
Ramin, P., Brock, A.L., Polesel, F., Causanilles, A., Emke, E., de Voogt, P., Plósz, B.G., 2016. Transformation and sorption of illicit drug biomarkers in sewer systems: understanding the role of suspended solids in raw wastewater. Environ. Sci. Technol. 50, 13397–13408.
Sin, G., Gernaey, K. V, Lantz, A.E., 2009. Good modelling practice (GMoP) for PAT applications: Propagation of input uncertainty and sensitivity analysis. Biotechnol. Prog. 25, 1043–1053.
Sin, G., Van Hulle, S.W.H., De Pauw, D.J.W., Van Griensven, A., Vanrolleghem, P.A., 2005. A critical comparison of systematic

calibration protocols for activated sludge models: A SWOT analysis. Water Res. 39, 2459–2474.

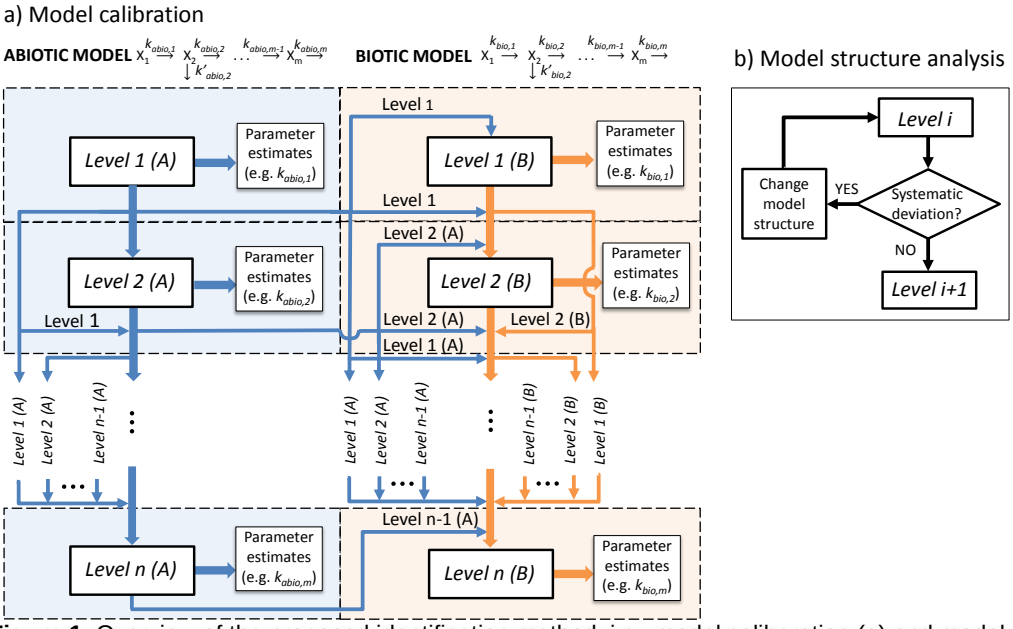

Figure 1. Overview of the proposed identification method, i.e., model calibaration (a) and model structure analysis (b), referred to as *Method 1* in this study, to estimate transformation rate constants in metabolic pathway models for both abiotic and biotic processes. The method includes *n* calibration levels for *m* number of metabolites (e.g., X_1 to X_m).

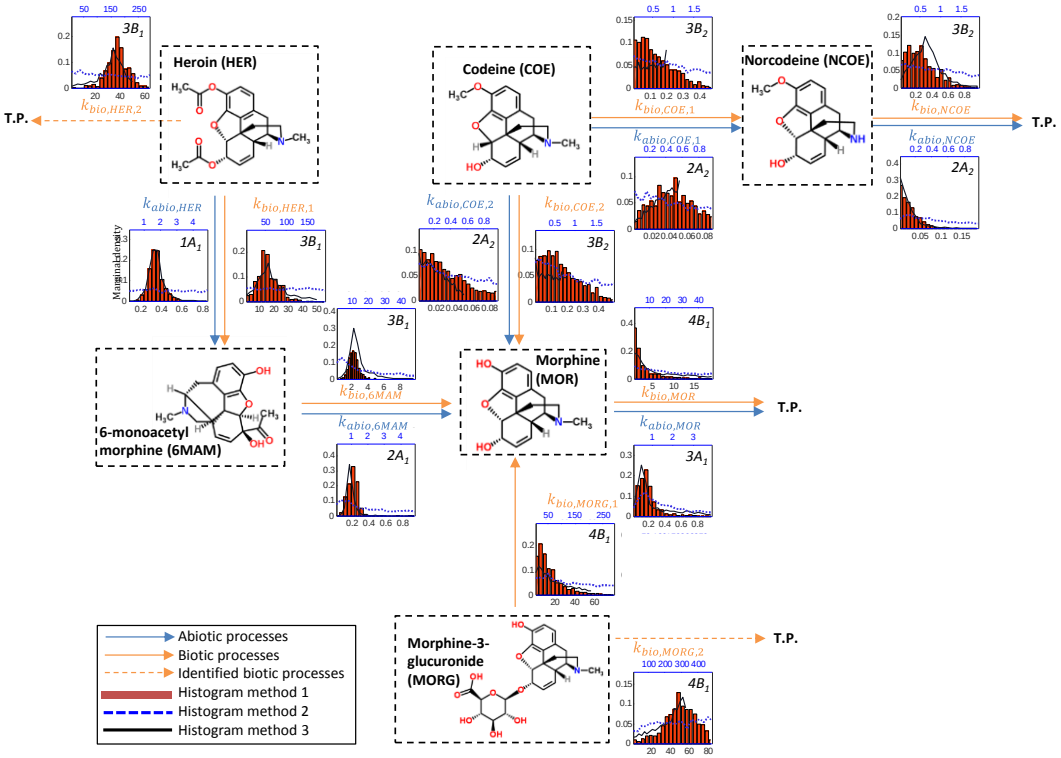

Figure 2. The identified transformation pathway of heroin (HER) and codeine (COE) drug biomarkers, including abiotic and biotic processes. Posterior distribution of estimated parameters (histograms) for the proposed methodology – *Method 1* (in red), *Method 2* (dotted blue line, upper X axis) and *Method 3* (solid black line).

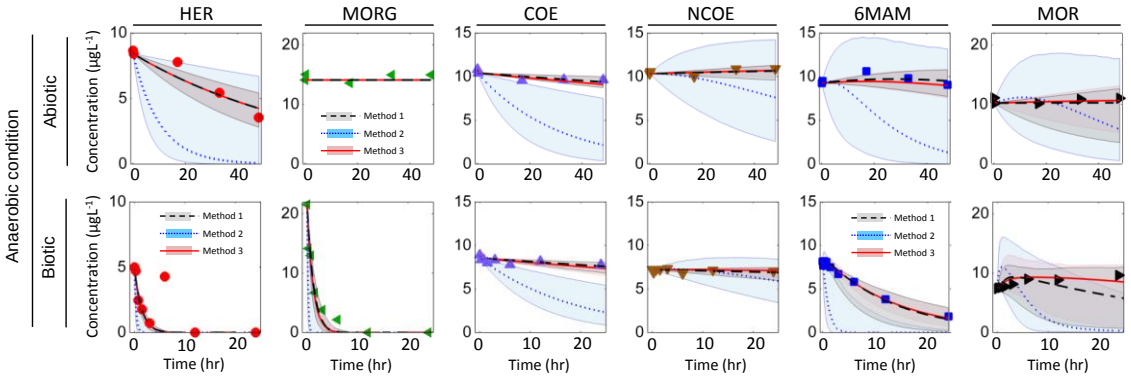

Figure 3. Measured and simulated biomarker concentration data with uncertainty bands obtained using Method 1-3. Markers are measured data and lines are simulation results. The shaded area reflects 95% credibility interval of model prediction (red area and full line: *Method 1*; gray area and dashed line: *Method 2*; blue area and dotted line: *Method 3*).

Table 1. Assessment of simulation accuracy of models calibrated using Methods 1-3. Abbreviations: RMSE, root mean squared error; MAE, mean absolute error; ILTC, interval length to coverage. For better comparison among the three methods, the following color code is used: green – high accuracy; yellow - moderate accuracy; red – low accuracy.

		HER			MORG			COE			NCOE			6MAM			MOR		
		Method 1	Method 2	Method 3	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3	Method 1	Method 2	Method 3
Abiotic model	RMSE	0.62	3.91	0.62				0.31	4.60	0.25	0.28	1.74	0.29	0.58	4.93	0.49	0.29	2.58	0.46
	MAE	0.42	2.98	0.42	na.			0.24	3.44	0.18	0.23	1.14	0.25	0.40	3.74	0.35	0.26	1.74	0.36
	ILTC	0.45	1.05	0.45				0.08	1.08	0.08	0.08	0.66	0.12	0.23	0.83	0.23	0.40	1.16	0.49
Biotic model	RMSE	0.25	1.25	0.25	1.58	5.72	1.54	0.29	2.25	0.27	0.19	0.40	0.18	0.23	3.89	0.23	0.57	4.52	1.41
	MAE	0.14	0.78	0.13	1.05	3.74	0.93	0.23	1.34	0.20	0.12	0.24	0.12	0.17	3.02	0.19	0.43	3.38	0.87
	ILTC	0.30	0.95	0.28	0.54	1.00	0.45	0.08	0.53	0.08	0.06	0.23	0.08	0.18	0.90	0.25	0.57	1.02	0.54

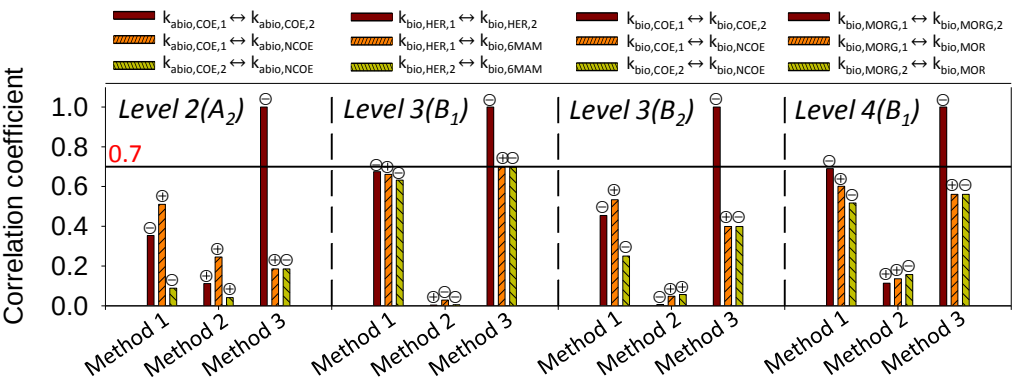

Figure 4. Linear correlation coefficients for model parameters following parameter estimation in calibration *Method 1-3*. Correlation threshold for identifiability defined at 0.7 according to Sin et al. (2009). Positive and negative correlations are designated with + and – signs, respectively.